

УДК 517.917

ББК 22.1

АЛЬТРУИСТИЧЕСКОЕ ПОВЕДЕНИЕ В НЕАНТАГОНИСТИЧЕСКОЙ ПОЗИЦИОННОЙ ДИФФЕРЕНЦИАЛЬНОЙ ИГРЕ

АНАТОЛИЙ Ф. КЛЕЙМЕНОВ*

Институт математики и механики

им. Н.Н. Красовского УрО РАН

620990, Екатеринбург, ул. С. Ковалевской, 16

e-mail: kleimenov@imm.uran.ru

Рассматривается неантагонистическая позиционная дифференциальная игра (НПДИ) двух лиц с динамикой, описываемой обыкновенным нелинейным векторным дифференциальным уравнением. Ограничения на значения управлений игроков являются геометрическими. Момент окончания игры фиксирован. Функционалы выигрыша обоих игроков являются терминальными. Формализация позиционных стратегий в НПДИ основана на формализации и результатах общей теории антагонистических позиционных дифференциальных игр (АПДИ) (см. монографии Н.Н. Красовского и А.И. Субботина [3, 4]). В настоящей статье дополнительно предполагается, что каждый игрок помимо обычного, *нормального* (*nor*), типа поведения, ориентированного на максимизацию собственного функционала, может использовать другие типы поведения, введенные в [1, 2]. В частности, это могут быть *альтруистический* (*alt*),

©2015 А.Ф. Клейменов

* Работа выполнена в рамках Программы Президиума РАН «Математические задачи современной теории управления».

агрессивный (*agg*) и *парадоксальный* (*par*) типы. Далее полагается, что по ходу игры игроки могут осуществлять переключения своего поведения с одного типа на другой. Использование игроками возможности такого переключения в повторяющейся биматричной 2×2 игре позволило в работах [5, 6] получить новые решения этой игры. В настоящей статье распространение указанного подхода на неантагонистические позиционные дифференциальные игры приводит к новым постановкам задач. В частности, представляет интерес как трансформируются выигрыши игроков, получаемые на нэшевских решениях. Актуальной становится задача минимизации времени «ненормального» поведения при условии достижения удовлетворительного результата. В статье предлагается формализация НПДИ двух лиц с типами поведения (НПДИсТП). Предполагается, что в НПДИсТП каждый игрок одновременно с выбором позиционной стратегии выбирает также индикаторную функцию, определенную на всем отрезке игры и принимающую значения в множестве $\{nor, alt, agg, par\}$. Индикаторная функция игрока показывает динамику изменения типа поведения, которой придерживается этот игрок. Таким образом, в НПДИсТП каждый игрок управляет выбором пары действий {позиционная стратегия, индикаторная функция}. Дается определение понятия *BT*-решения такой игры. Ожидаемо, что в НПДИсТП использование типов поведения, отличных от нормального (так называемых *ненормальных* (*abnormal*) типов), в ряде случаев может привести к исходам, более предпочтительным для игроков, чем в игре НПДИ. В статье рассматриваются два примера игры с динамикой простого движения на плоскости, в каждом из которых один игрок придерживается альтруистического типа поведения в течение некоторого промежутка времени. Показывается, что по сравнению с игрой с нормальными типами поведения игроков, в первом примере на *BT*-решении происходит увеличение выигрыша каждого из игроков, а во втором примере – увеличивается суммарный выигрыш игроков.

Ключевые слова: неантагонистическая позиционная дифференциальная игра, терминальные показатели качества, типы поведения игроков, альтруистический тип поведения, решения нэшевского типа.

1. Уравнения движения, функционалы, стратегии и движения

Динамика неантагонистической позиционной дифференциальной игры (НПДИ) описывается уравнением

$$\dot{x} = f(t, x, u, v), \quad t \in [t_0, \vartheta], \quad x(t_0) = x_0, \quad (1.1)$$

где $x \in R^n$ – фазовый вектор; управления первого и второго игроков u и v стеснены ограничениями $u \in P \in \text{comp}R^p$ и $v \in Q \in \text{comp}R^q$. В пространстве переменных t, x задано компактное множество G , проекция которого на ось t равна отрезку $[t_0, \vartheta]$, где ϑ – заданный момент окончания игры. Предполагается, что функция $f : G \times P \times Q \rightarrow R^n$ непрерывна, липшицева по x и удовлетворяет условию подлинейного роста по x . Кроме того, полагаем, что выполнено условие седловой точки в маленькой игре ([4], стр.56).

Функционалы выигрыша игроков имеют вид

$$I_i = \sigma_i(x(\vartheta)), \quad i = 1, 2, \quad (1.2)$$

где $\sigma_i(\cdot)$ – непрерывные функции.

Предполагается, что оба игрока имеют полную информацию о текущей позиции (t, x) игры. Используемая в работе формализация НПДИ, включающая описание позиционных стратегий игроков, а также порождаемых этими стратегиями конструктивных и предельных движений, основана на формализации и результатах [3, 4] и подробно описана в [1]. Напомним кратко основные элементы этой формализации.

Позиционная стратегия игрока 1 отождествляется с парой $U = \{u(t, x, \varepsilon), \beta_1(\varepsilon)\}$, где $u(\cdot)$ – произвольная функция позиции (t, x) и положительного параметра точности ε , принимающая значения во множестве P . Функция $\beta_1 : (0, \infty) \rightarrow (0, \infty)$ непрерывна, монотонна и удовлетворяет условию $\beta_1(\varepsilon) \rightarrow 0$ при $\varepsilon \rightarrow 0$. При фиксированном ε величина $\beta_1(\varepsilon)$ задает верхнюю границу для шага разбиения отрезка $[t_0, \vartheta]$, используемого игроком 1 при построении аппроксимационных движений. Аналогично определяется позиционная стратегия игрока 2 – $V = \{v(t, x, \varepsilon), \beta_2(\varepsilon)\}$.

Рассматриваются два типа движений, порождаемых парой стратегий (U, V) из начальной позиции (t_0, x_0) : аппроксимационные (ло-

маные Эйлера) и идеальные (предельные) движения. Аппроксимационное движение $x_{\Delta}^{\varepsilon}[t] = x[t, t_0, x_0, U, \varepsilon_1, \Delta_1, V, \varepsilon_2, \Delta_2]$ определяется при выбранных игроками 1 и 2 значениях параметров точности ε_1 и ε_2 , а также разбиениях $\Delta_1 = \{t_i^{(1)}\}$ и $\Delta_2 = \{t_k^{(2)}\}$ отрезка $[t_0, \vartheta]$. При этом игрок i , $i = 1, 2$ в качестве своего управления использует кусочно-постоянную функцию со значениями, вычисляемыми в узлах своего разбиения, для шага которого выполнено условие $\delta(\Delta_i) \leq \beta_i(\varepsilon_i)$, где $\delta(\Delta_i) = \max_j (t_{j+1}^{(i)} - t_j^{(i)})$. Идеальное (предельное) движение $x(t) = x(t, t_0, x_0, U, V)$ определяется как равномерный предел последовательности аппроксимационных движений

$$\{x_{\Delta^s}^{\varepsilon^s}[t, t_0^s, x_0^s, U, \varepsilon_1^s, \Delta_1^s, V, \varepsilon_2^s, \Delta_2^s]\}$$

при $s \rightarrow \infty$, $\varepsilon_i^s \rightarrow 0$, $t_0^s \rightarrow t_0$, $x_0^s \rightarrow x_0$, $\delta(\Delta_i^s) \leq \beta_i(\varepsilon_i^s)$, $i = 1, 2$.

Множество предельных движений $X(t_0, x_0, U, V)$ является компактным в метрике пространства $C[t_0, \vartheta]$.

2. Некоторые результаты из теории НПДИ

В дальнейшем нам потребуются некоторые результаты из теории НПДИ [1].

Определение 2.1. *Пара стратегий (U^N, V^N) образует равновесное по Нэшу решение (NE-решение) в НПДИ (1.1)–(1.2), если для любого движения $\bar{x}(\cdot) \in X(t_0, x_0, U^N, V^N)$, любого $\tau \in [t_0, \vartheta]$, любых стратегий U и V имеют место следующие неравенства:*

$$\max_{x(\cdot)} \sigma_1(x[\vartheta, \tau, \bar{x}(\tau), U, V^N]) \leq \min_{x(\cdot)} \sigma_1(x[\vartheta, \tau, \bar{x}(\tau), U^N, V^N]), \quad (2.1)$$

$$\max_{x(\cdot)} \sigma_2(x[\vartheta, \tau, \bar{x}(\tau), U^N, V]) \leq \min_{x(\cdot)} \sigma_2(x[\vartheta, \tau, \bar{x}(\tau), U^N, V^N]), \quad (2.2)$$

где операции \max и \min производятся по соответствующим множествам предельных движений.

Определение 2.2. *NE-решение (U^P, V^P) , неумлучшаемое по Парето относительно величин (I_1, I_2) (1.2), называется P^* -решением.*

Рассмотрим вспомогательные антагонистические позиционные дифференциальные игры Γ_1 и Γ_2 . Динамика обеих игр описывается уравнением (1.1). В игре Γ_i игрок i максимизирует функционал

$\sigma_i(x(\vartheta))$ (1.2), а игрок $3 - i$ ему противодействует. Из [3] следует, что обе игры Γ_1 и Γ_2 имеют универсальные седловые точки

$$u^{(i)}(t, x, \varepsilon), v^{(i)}(t, x, \varepsilon), \quad i = 1, 2 \quad (2.3)$$

и непрерывные функции цены $\gamma_1(t, x)$ и $\gamma_2(t, x)$. Свойство универсальности стратегий (2.3) означает, что они являются оптимальными не только для фиксированной начальной позиции (t_0, x_0) , но и для любой позиции $(t_*, x_*) \in G$, рассматриваемой в качестве начальной. Очевидно, что для позиции $(t, x) \in G$ игры НПДИ величина $\gamma_i(t, x)$ представляет собою гарантированный выигрыш игрока i в этой позиции.

В [1] показано, что все NE - и P^* -решения игры могут быть найдены в классе пар стратегий (U, V) , порождающих единственное предельное движение (траекторию). Решающие стратегии, составляющие такую пару, которая порождает траекторию $x^*(\cdot)$, имеют следующий вид:

$$U^0 \div \{u^0(t, x, \varepsilon), \beta_1^0(\varepsilon)\}, \quad V^0 \div \{v^0(t, x, \varepsilon), \beta_2^0(\varepsilon)\}, \quad (2.4)$$

$$u^0(t, x, \varepsilon) = \begin{cases} u^*(t, \varepsilon), & \text{если } \|x - x^*(t)\| < \varepsilon\varphi(t), \\ u^{(2)}(t, x, \varepsilon), & \text{если } \|x - x^*(t)\| \geq \varepsilon\varphi(t), \end{cases} \quad (2.5)$$

$$v^0(t, x, \varepsilon) = \begin{cases} u^*(t, \varepsilon), & \text{если } \|x - x^*(t)\| < \varepsilon\varphi(t), \\ v^{(1)}(t, x, \varepsilon), & \text{если } \|x - x^*(t)\| \geq \varepsilon\varphi(t) \end{cases}$$

для всех $t \in [t_0, \vartheta], \varepsilon > 0$. В (2.4)–(2.5) через $u^*(t, \varepsilon)$, $v^*(t, \varepsilon)$ обозначены семейства программных управлений, порождающих траекторию $x^*(t)$. Функция $\varphi(\cdot)$ вместе с функциями $\beta_1^0(\cdot)$ и $\beta_2^0(\cdot)$ выбираются таким образом, что ломаные Эйлера, порожденные парой (U^0, V^0) из начальной позиции (t_0, x_0) , не выходят за пределы $\varepsilon\varphi(t)$ -окрестности траектории $x^*(t)$. Функции $u^{(2)}(\cdot, \cdot, \cdot)$ и $v^{(1)}(\cdot, \cdot, \cdot)$ определены в (2.3).

Далее, для каждой NE - и P^* -траектории $x^*(t)$ имеет место следующее свойство: точка $t = \vartheta$ является точкой максимума функции гарантированного выигрыша игрока i , вычисленной вдоль этой траектории, т.е.

$$\max_{t \in [t_0, \vartheta]} \gamma_i(t, x^*(t)) = \gamma_i(\vartheta, x^*(\vartheta)), \quad i = 1, 2. \quad (2.6)$$

3. Типы поведения игроков

Дополнительно предполагаем, что помимо обычного, *нормального* (*nor*), типа поведения, ориентированного на максимизацию собственных функционалов выигрыша (1.2), игроки могут использовать другие типы поведения, например, введенные в работах [2, 5]. А именно, это следующие типы: *альтруистический* («чем лучше моему партнеру, тем лучше мне»), *агрессивный* («чем хуже моему партнеру, тем лучше мне») и *парадоксальный* («чем хуже мне, тем лучше мне»).

Эти три типа поведения могут быть формализованы следующим образом.

Определение 3.1. Скажем, что игрок 1 придерживается на отрезке $[t_*, t^*]$ альтруистического (*alt*) типа поведения по отношению к игроку 2, если на этом отрезке действия игрока 1 направлены на максимизацию функционала I_2 игрока 2.

Определение 3.2. Скажем, что игрок 1 придерживается на отрезке $[t_*, t^*]$ агрессивного (*agg*) типа поведения по отношению к игроку 2, если на этом отрезке действия игрока 1 направлены на минимизацию функционала I_2 игрока 2.

Определение 3.3. Скажем, что игрок 1 придерживается на отрезке $[t_*, t^*]$ парадоксального (*par*) типа поведения, если на этом отрезке действия игрока 1 направлены на минимизацию собственного функционала I_1 .

Аналогично определяются альтруистический и агрессивный типы поведения игрока 2 по отношению к игроку 1, а также парадоксальный тип поведения игрока 2. Заметим, что агрессивный тип поведения игроков фактически используется в НПДИ в форме стратегий наказания, содержащихся в структуре решений игры (см., например, [1]).

Приведенные определения характеризуют экстремальные типы поведения игроков. В действительности же реальные индивидуумы ведут себя, как правило, частично нормально, частично альтруистично, частично агрессивно и частично парадоксально. Другими словами, смешанные типы поведения, по-видимому, больше согласуются с реальностью.

Если каждого игрока ограничить «чистыми» типами поведения, то в рассматриваемой игре двух лиц (1.1) с функционалами I_1 и I_2 (1.2) существует 16 возможных комбинаций, показанных в табл. 1. При этом в четырех комбинациях интересы игроков совпадают, и игроки решают командные задачи управления (team problems). В других четырех комбинациях игроки имеют противоположные интересы и, следовательно, разыгрываются антагонистические дифференциальные игры. Остальные 8 пар определяют неантагонистические дифференциальные игры.

Таблица 1.

	<i>nor</i>	<i>alt</i>	<i>agg</i>	<i>par</i>
<i>nor</i>	(I_1, I_2)	(I_1, I_1)	$(I_1, -I_1)$	$(I_1, -I_2)$
<i>alt</i>	(I_2, I_2)	(I_2, I_1)	$(I_2, -I_1)$	$(I_2, -I_2)$
<i>agg</i>	$(-I_2, I_2)$	$(-I_2, I_1)$	$(-I_2, -I_1)$	$(-I_2, -I_2)$
<i>par</i>	$(-I_1, I_2)$	$(-I_1, I_1)$	$(-I_1, -I_1)$	$(-I_1, -I_2)$

Идея использования игроками возможности переключения по ходу игры своего поведения с одного типа на другой была применена для игры с кооперативной динамикой в работе [5] и для повторяющейся биматричной 2×2 игры в работе [6], что позволило получить новые решения в этих играх.

Распространение указанного подхода на неантагонистические позиционные дифференциальные игры приводит к новым постановкам задач. В частности, представляет интерес как трансформируются выигрыши игроков, получаемые на нэшевских решениях. Актуальной становится задача минимизации времени «ненормального» поведения при условии достижения результата, более хорошего, чем при нормальном поведении игроков.

Далее будем полагать, что одновременно с выбором позиционной стратегии каждый игрок выбирает также свою индикаторную функцию, определенную на отрезке $t \in [t_0, \vartheta]$ и принимающую значение в множестве $\{nor, alt, agg, par\}$. Обозначим индикаторную функцию игрока i символом $\alpha_i : [t_0, \vartheta] \rightarrow \{nor, alt, agg, par\}$, $i = 1, 2$. Если индикаторная функция какого-то игрока принимает значение, скажем, *alt* на некотором отрезке времени, то этот игрок действует на этом отрезке как альтруист по отношению к своему партнеру по игре.

Таким образом, в рассматриваемой игре с различными типами поведения игроков игрок 1 управляет выбором пары *действий* {позиционная стратегия, индикаторная функция}: $(U, \alpha_1(\cdot))$, а игрок 2 управляет выбором пары действий $(V, \alpha_2(\cdot))$. Далее такую НПДИ с типами поведения будем обозначать через НПДИсТП.

Заметим, что если индикаторные функции обоих игроков тождественно равны значению *por* на всем отрезке игры, то имеем классическую НПДИ.

4. BT-решение игры НПДИсТП

Рассмотрим теперь НПДИсТП с классами действий игроков 1 и 2:

$$(U, \alpha_1(\cdot)), (V, \alpha_2(\cdot)). \quad (4.1)$$

Очевидно, что множество движений, порожденных парой действий (4.1), совпадает с множество движений, порожденных парой (U, V) в соответствующей НПДИ.

Определение 4.1. Пара $\{(U^0, \alpha_1^0(\cdot)), (V^0, \alpha_2^0(\cdot))\}$ образует сильное BT-решение игры НПДИсТП, если найдется порожденная парой траектория $x^{BT}(\cdot)$ и найдется P^* -решение в соответствующей игре НПДИ, порождающее траекторию $x^P(\cdot)$, такие, что

$$\sigma_i(x^{BT}(\vartheta)) > \sigma_i(x^P(\vartheta)), \quad i = 1, 2. \quad (4.2)$$

Определение 4.2. Пара $\{(U^0, \alpha_1^0(\cdot)), (V^0, \alpha_2^0(\cdot))\}$ образует слабое BT-решение игры НПДИсТП, если найдется порожденная парой траектория $x^{BT}(\cdot)$ и найдется P^* -решение в соответствующей игре НПДИ, порождающее траекторию $x^P(\cdot)$, такие, что

$$\sum \sigma_i(x^{BT}(\vartheta)) > \sum \sigma_i(x^P(\vartheta)). \quad (4.3)$$

Очевидно, что сильное BT-решение является слабым BT-решением; обратное, вообще говоря, неверно.

Задача 4.1. Найти множество (сильных и слабых) BT-решений.

Задача 4.2. Найти множество BT-решений, доставляющих минимум времени использования игроками *ненормальных* типов поведения.

Вполне ожидаемо, что использование игроками в игре НПДИсТП типов поведения, отличных от нормального, может в ряде случаев привести к исходам, более предпочтительным для них, чем в соответствующей игре НПДИ только с нормальным типом поведения.

Далее приводятся два примера игры с динамикой простого движения на плоскости, в каждом из которых один из игроков в течение некоторого промежутка времени придерживается альтруистического типа поведения. Показывается, что по сравнению с игрой с нормальным типом поведения обоих игроков, в первом примере происходит увеличение выигрыша каждого игрока, а во втором примере игрок-альтруист значительно увеличивает выигрыш другого игрока, уменьшая несколько свой выигрыш; при этом суммарный выигрыш игроков увеличивается. В обоих примерах решается задача нахождения *BT*-решений, доставляющих минимум времени ненормального поведения.

5. Пример 1

Пусть динамика (1.1) имеет вид

$$\dot{x} = u + v, \quad x, u, v \in R^2, \quad \|u\| \leq 1, \quad \|v\| \leq 1, \quad 0 \leq t \leq \vartheta, \quad x(0) = x_0, \quad (5.1)$$

а функционалы выигрыша (1.2) суть

$$I_i = 10 - \|(x(\vartheta) - a^{(i)})\|, \quad i = 1, 2, \quad (5.2)$$

т.е. игрок i стремится привести точку $x(\vartheta)$ как можно ближе к своей целевой точке $a^{(i)}$. Зададим следующие значения параметров игры: $\vartheta = 2.5$, $x_0 = (0, 0)$, $a^{(1)} = (7.0, 5.0)$, $a^{(2)} = (-7.0, 5.0)$.

Примем еще дополнительное условие. В плоскости (x_1, x_2) круг S радиуса 2 с центром в точке $m = (0, 2.5)$ задает следующее фазовое ограничение в задаче: траекториям системы (5.1) запрещается заходить во внутренность круга S .

На рис. 1 кривые $\bar{d}a$ и ad вместе с дугой $\bar{a}сса$ границы множества S и дугой $\bar{d}qd$ окружности радиуса 5 ограничивают множество достижимости системы (5.1) с введенными фазовыми ограничениями, построенное для момента $\vartheta = 2.5$. Линия $Ос$ является касательной к окружности S . Точка b является точкой круга S , ближайшей

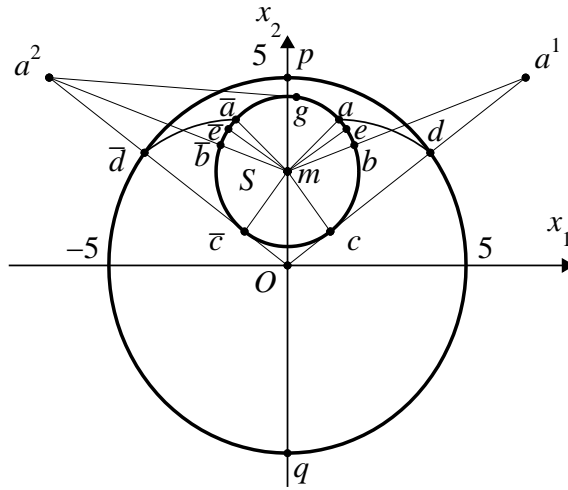


Рисунок 1. К примеру 1

к $a^{(1)}$. Результаты вычислений (здесь и далее указываются приближенные числовые значения): $a = (1.36, 3.96)$, $c = (1.20, 0.90)$, $b = (1.88, 3.17)$. Рис. 1 обладает симметрией относительно оси ординат. Обозначения симметричных точек отличаются только крышками. Дальнейшее описание производится в основном для правой половины рисунка.

В игре НПДИ с нормальным типом поведения игроков нэшевской (и одновременно P^* -траекторией) будет траектория $x(t) \equiv 0$, $t \in [0, 2.5]$ (стационарная точка O), на которой игроки получают выигрыши $I_1 = I_2 = 10 - \sqrt{74} \approx 1.40$. Траектория, изображаемая линией $Ocba$, не является нэшевской, поскольку в точке b игрок 1 получает максимальный выигрыш (равный 4.57) и не будет заинтересован в дальнейшем отслеживании траектории до точки a . Траектория, изображаемая линией Ocb , также не является нэшевской: теперь уже игрок 2 не согласится на отслеживание траектории, поскольку в точке b он получит выигрыш, равный 0.53, что меньше, чем в точке O .

Переходим теперь к игре НПДИсТП и рассмотрим случай, когда игрок 1 проявляет альтруизм по отношению к игроку 2 в течение некоторого промежутка времени.

На окружности S находим точку e , расстояние от которой до точки $a^{(2)}$, вычисленное вдоль линии $ega^{(2)}$ ($a^{(2)}g$ – касательная к окружности S), равно длине отрезка $Oa^{(2)}$. Получаем $e = (1.43, 3.90)$. Время попадания в точку e при движении с максимальной скоростью равно $t = 2.45$. Если в момент $t = 2.45$ продолжить движение вдоль дуги ea против часовой стрелки, то расстояние до точки $a^{(2)}$ станет меньшим, чем длина отрезка $Oa^{(2)}$. Поэтому игроку 2 будет выгодно завершить движение в какой-либо точке дуги ea .

Пусть h – некоторая фиксированная точка дуги ea , а t^* – время попадания в точку h при движении с максимальной скоростью. Очевидно, что $t^* \in [2.45, 2.5]$.

Рассмотрим пару стратегий (U^*, V^*) , порождающую единственное предельное движение, изображаемое линией $Oceh$ при $t \in [0, t^*]$, и остающееся стационарным в точке h при $t \in [t^*, 2.5]$. Далее, рассмотрим следующие индикаторные функции игроков $\alpha_1^*(t) = \{nor, t \in [0, 2.01]; alt, t \in [2.01, t^*]; nor, t \in [t^*, 2.5]\}$ и $\alpha_2^*(t) = \{nor, t \in [0, 2.5]\}$. Здесь $t = 2.01$ – момент попадания в точку b при движении с максимальной скоростью.

Нетрудно видеть, что оба игрока заинтересованы в реализации стратегий (U^*, V^*) и индикаторных функций-программ $\alpha_1^*(\cdot), \alpha_2^*(\cdot)$.

Таким образом, получаем зависящее от параметра $t^* \in [2.45, 2.5]$ однопараметрическое семейство наборов действий игроков

$$\{(U^*, \alpha_1^*(\cdot)), (V^*, \alpha_2^*(\cdot))\}, \quad (5.3)$$

каждый из которых является BT -решением.

Если значение параметра $t^* = 2.45$, то имеем слабое BT -решение; при остальных значениях $t^* \in (2.45, 2.5]$ каждый из наборов (5.3) является сильным BT -решением.

Если построенная таким образом BT -траектория заканчивается в точке a (это будет при $t^* = 2.5$), то выигрыши игроков будут $I_1 = 4.2$, $I_2 = 1.50$, то есть оба игрока выигрывают по сравнению с игрой с нормальным типом поведения.

Решение задачи 4.2 достигается на наборе (5.3) при значении $t^* = 2.45$.

6. Пример 2

Пусть динамика (1.1) имеет вид

$$\dot{x} = u + v, \quad x, u, v \in R^2, \quad \|u\| \leq 1, \quad \|v\| \leq 1, \quad 0 \leq t \leq \vartheta, \quad x(0) = x_0, \tag{6.1}$$

а функционалы выигрыша (1.2) имеют вид

$$I_1 = \sigma_1(x(\vartheta)) = 9 - \|x(\vartheta) - a^{(1)}\|, \tag{6.2}$$

$$I_2 = \sigma_2(x(\vartheta)) = 5 + \sqrt{3}|x_1(\vartheta)| - x_2(\vartheta), \tag{6.3}$$

т.е. цель игрока 1 – привести вектор $x(\vartheta)$ как можно ближе к целевой точке $a^{(1)}$.

Зададим следующие значения параметров игры: $\vartheta = 1.5$; $x_0 = (-2, 2\sqrt{3})$; $a^{(1)} = (5, 7.5)$. На рис. 2 в плоскости (x_1, x_2) круг радиуса 3 с центром в начальной точке $A(-2, 2\sqrt{3})$ изображает множество достижимости системы (6.1), построенное для момента $\vartheta = 1.5$. Начальная точка A лежит на линии уровня функции $\sigma_2(x_1, x_2) = 5 + \sqrt{3}|x_1| - x_2 = 5$.

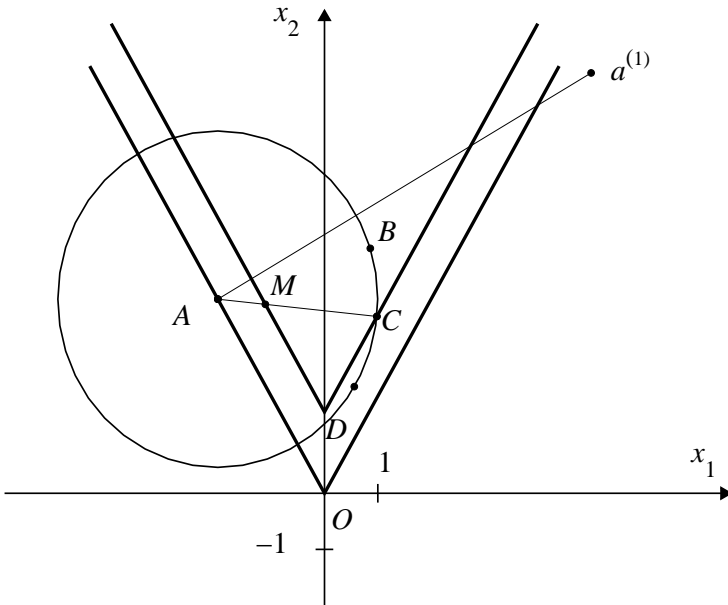


Рисунок 2. К примеру 2

Поскольку $AO \perp Aa^{(1)}$, то в игре НПДИ с нормальным типом поведения игроков единственной нэшевской траекторией, а, следовательно, и единственной траекторией, порожденной P^* -решением, будет траектория $x(t) = 0$, $t \in [0, 1.5]$, которая совпадает со стационарной точкой A . Вычисления показывают, что выигрыши игроков на этом P^* -решении равны $I_1 = 0.920$, $I_2 = 5.000$; сумма выигрышей игроков равна $I_1 + I_2 = 5.920$ (здесь и ниже указаны приближенные значения).

Переходя к игре НПДИСТП, прежде всего отметим, что в рассматриваемом примере заведомо исключена возможность существования сильных BT -решений, поскольку во множестве достижимости нет точек, в которых оба игрока получают выигрыши, большие, чем на P^* -решении. Действительно, пересечение двух подмножеств множества достижимости, каждое из которых состоит из точек, в которых выигрыш одного из игроков больше, чем в точке A , пусто.

Будем искать слабые BT -решения и рассмотрим случай, когда игрок 2 проявляет альтруизм по отношению к игроку 1 в течение некоторого промежутка времени. Во множестве достижимости находим все точки, в которых сумма выигрышей игроков $I_1 + I_2$ больше или равна сумме выигрышей на P^* -решении. Вычисления показывают, что множество T таких точек непусто и пересечением его с границей множества достижимости является дуга BD , где $B(0.86, 4.37)$, $D(0.56, 1.90)$. При этом точки дуги BD доминируют по сумме выигрышей остальные точки множества T . Далее, находим точку $C(0.984, 3.155)$, доставляющую максимум сумме выигрышей игроков на дуге BD . Проходящая через точку C линия уровня функции $\sigma_2(x_1, x_2) = 5 + \sqrt{3}|x_1| - x_2 = 3.549$ пересекает отрезок AC еще в одной точке $M(-1.109, 3.371)$.

Нетрудно видеть, что отрезок AC будет представлять траекторию, порожденную слабым BT -решением игры, если при движении с максимальной скоростью от начальной точки A до точки M (на этой части траектории игрок 2 уменьшает свой гарантированный выигрыш со значения 5 до значения 3.083) игрок 2 проявляет альтруизм по отношению к игроку 1, а после прохождения точки M игрок 2 возвращается к нормальному типу поведения. Поскольку расстояние между точками A и M равно 0.896, то продолжительность альтруистического поведения игрока 2 будет равна $0.5 \cdot 0.896 = 0.448$.

Рассмотрим пару стратегий (U^*, V^*) , порождающую при $t \in [0, 1.5]$ единственное предельное движение, изображаемое отрезком AC . И рассмотрим следующие индикаторные функции игроков $\alpha_1^*(t) = \{nor, t \in [0, 1.5]\}$ и $\alpha_2^*(t) = \{alt, t \in [0, 0.448]; nor, t \in [0.448, 1.5]\}$. Нетрудно видеть, что оба игрока заинтересованы в реализации стратегий (U^*, V^*) и индикаторных функций-программ $\alpha_1^*(\cdot), \alpha_2^*(\cdot)$.

Таким образом, набор действий игроков $(U^*, \alpha_1^*(\cdot)), (V^*, \alpha_2^*(\cdot))$, порождающий единственную траекторию AC , представляет собою слабое BT -решение. На этом BT -решении выигрыши игроков равны $I_1 = 3.083$, $I_2 = 3.549$; сумма выигрышей игроков равна $I_1 + I_2 = 6.632$. То есть по сравнению с игрой с нормальным типом поведения игрок 2 несколько проиграл, а игрок 1 значительно выиграл, так что суммарный выигрыш игроков увеличился.

Можно построить и другие слабые BT -решения, которые порождают единственные предельные движения, совпадающие с отрезками, соединяющими точку A с точками дуги BD , исключая ее концевые точки. Соответствующие наборы действий игроков строятся вышеописанным образом.

СПИСОК ЛИТЕРАТУРЫ

1. Клейменов А.Ф. *Неантагонистические позиционные дифференциальные игры*. Екатеринбург: Наука, 1993.
2. Клейменов А.Ф. *О решениях в неантагонистической позиционной дифференциальной игре* // Прикладная математика и механика. 1997. Т. 61, вып. 5. С. 739–746.
3. Красовский Н.Н. *Управление динамической системой*. М.: Наука, 1985.
4. Красовский Н.Н., Субботин А.И. *Позиционные дифференциальные игры*. М.: Наука, 1974.
5. Kleimenov A.F., Kryazhimskii A.V. *Normal Behavior, Altruism and Aggression in Cooperative Game Dynamics*. Interim Report IR-98-076, Laxenburg: IIASA, 1998.

6. Kleimenov A.F. *An Approach to Building Dynamics for Repeated Bimatrix 2×2 Games Involving Various Behavior Types* // In: *Dynamic and Control* (Leitman G., ed), London: Gordon and Breach, 1998. P. 195–204.

ALTRUISTIC BEHAVIOR IN A NON-ANTAGONISTIC POSITIONAL DIFFERENTIAL GAME

Anatolii F. Kleimenov, IMM UrO RAN, Dr.Sc., leading researcher (kleimenov@imm.uran.ru).

Abstract: A two-person non-antagonistic positional differential game (NPDG) whose dynamics is described by an ordinary nonlinear vector differential equation is considered. Constraints on values of players' controls are geometrical ones. The final time of the game is fixed. Cost functionals of both players are terminal ones. The formalization of positional strategies in the NPDG is based on the formalization and the results of the general theory of antagonistic positional differential games (APDGs) (see monographs by N.N. Krasovskii and A.I. Subbotin [3, 4]). Additionally, in the present paper we assume that each player together with the usual, *normal* type of behavior, oriented the maximization of their own functional, can use other types of behavior, introduced in the works [2, 5]. In particular, it may be *altruistic*, *aggressive* and *paradoxical* types. Further it is assumed that in the course of the game players can shift their behavior from one type to another. Using the players' possibility of such switches in a repeated bimatrix 2×2 game allowed in the works [5, 6] to obtain new solutions of this game. In the present paper, the spread of this approach on the NPDG leads to a new formulation of the problem. In particular, it is interesting how the players' outcomes at Nash solutions are transformed. An urgent task is to minimize the time of abnormal behavior subject to achieve a good result.

The paper proposes a formalization of the NPDG with behavior types (NPDGwBT). It is assumed that in an NPDGwBT each player simultaneously with a choice of positional strategy chooses also his own function indicator defined on the whole interval of the game and taking values in the set $\{normal, altruistic, aggressive, paradoxical\}$. The indicator function shows the dynamics of changes in player's behavior type, which adheres to this player. Thus, in this NPDGwBT each player controls the selection of the pair {positional strategy, function indicator}. The definition of the concept of a BT-solution of such a game is given. Expectedly, that in NPDGwBT the using behavior types which differ from normal one (so-called *abnormal* types), in some cases, may lead to more favorable outcomes for the players than in the NPDG. In the paper two examples of the NPDGwBT with simple dynamics in the plane, in each of which one player keeps to altruistic type of behavior over a period of time, are considered. It is shown, that in the first example on BT-solution the payoffs of both players are increased, in comparison with the game with normal type of behavior, and in the second example the sum of players' payoffs is increased.

Keywords: non-antagonistic positional two-person differential game, terminal cost functionals, behavior types of players, altruistic type, solutions of Nash type.